INRIA

Team clime

Coupling environmental data and numerical simulation models for integrated software solution

Paris - Rocquencourt

Theme : Observation and Modeling for Environmental Sciences

Activity Report

2009

Table of contents

1. Team

Research Scientist

Jean-Paul Berroir [Research scientist, CR, INRIA, until August 2009]

Marc Bocquet [Deputy Director of CEREA, École des Ponts ParisTech, HdR]

Isabelle Herlin [Team Leader, Research Director, INRIA]

Vivien Mallet [Research scientist, CR, INRIA]

Faculty Member

Etienne Huot [MC, University of Versailles]

External Collaborator

Dominique Béréziat [MC, UPMC]

Technical Staff

Claire Mouton [Technical Staff, INRIA]

Nicolas Mercier [Technical Staff, INRIA]

Pierre Tran [Technical Staff, École des Ponts ParisTech]

PhD Student

Ahmed Abdel Meguid [PhD Student, INRIA, until June 2009]

Rachid Abida [PhD Student, École des Ponts ParisTech, until September 2009]

Karim Drifi [Internship, ENS Cachan, from May to November 2009,PhD Student, INRIA, from December 2009]

Damien Garaud [PhD Student, EdF R&D, École des Ponts ParisTech]

Irène Korsakissok [PhD Student, École des Ponts ParisTech]

Mohammad Reza Koohkan [PhD Student, École des Ponts ParisTech, from October 2009]

Post-Doctoral Fellow

Xiaoni Wang [Post-Doctoral Fellow, INRIA, until October 2009]

Lin Wu [Post-Doctoral Fellow, École des Ponts ParisTech]

Sergey Zhuk [Post-Doctoral Fellow, ERCIM, from October 2009]

Administrative Assistant

Véronique Dehlinger [Administrative Assistant, École des Ponts ParisTech]

Nathalie Gaudechoux [Administrative Assistant, INRIA]

Other

Imen Jdey [Internship, INRIA, from August to December 2009]

Sahar Syassi [Internship, INRIA, from March to August 2009]

2. Overall Objectives

2.1. Overall Objectives

The international political and scientific context is indicating the serious potential risks related to environmental problems and is pointing out the role that can be played by models and observation systems for the evaluation and forecasting of these risks. At the political level, agreements, such as the Kyoto protocol, European directives on air quality or on major accident hazards involving dangerous substances (Seveso directive), and the French Grenelle de l'Environnement establish objectives for the mitigation of environmental risks. These objectives are supported at a scientific level by international initiatives, like the European GMES program (Global Monitoring of Environment and Security), or national programs, such as the Air Chemistry program, which give a long term structure to environmental research. These initiatives emphasize the importance of observational data and the potential of satellite acquisitions.

The complexity of the environmental phenomena as well as the operational objectives of risk mitigation necessitate an intensive interweaving between physical models, data processing, simulation, visualization and database tools.

This situation is met for instance in atmospheric pollution, an environmental domain whose modeling is gaining an ever-increasing significance and impact, either at local (air quality), regional (transboundary pollution) or global scale (greenhouse effect). In this domain, modeling systems are used for operational forecasts (short or long term), detailed case studies, impact studies for industrial sites, as well as coupled modeling (e.g., pollution and health, pollution and economy). These scientific subjects strongly require linking the models with all available data either of physical origin (e.g., models outputs), coming from raw observations (satellite acquisitions and/or information measured *in situ* by an observation network) or obtained by processing and analysis of these observations (e.g., chemical concentrations retrieved by inversion of a radiative transfer model).

Clime has been jointly created, by INRIA and École des Ponts ParisTech, for studying these questions with researchers in data assimilation, image processing, and modeling.

Clime carries out research activities in three main area:

- Data assimilation methods: inverse modeling, network design, ensemble methods, uncertainties estimation, ...

- Image assimilation: assimilating structures within environmental forecasting models, solving ill-posed image processing problems by image assimilation, defining dynamic models from images.

- Development of integrated chains for data/models/outputs (system architecture, workflow, database, visualization, ...).

3. Scientific Foundations

3.1. Data assimilation and inverse modeling

This activity is currently one of the major concerns of environmental sciences. It matches up the setting and the use of data assimilation methods, for instance variational methods (4D-var). An emerging issue lies in the propagation of uncertainties in models, notably through ensemble forecasting methods.

Although modeling is not part of the scientific objectives of Clime, we have complete access to models developed by CEREA (Joint Laboratory of École des Ponts ParisTech/EDF R&D): the models from Polyphemus (pollution forecasting from local to regional scales) and Code-Saturne (urban scale). In regard to other modeling domains, Clime accesses models through co-operation initiatives either directly (for instance the shallow water model developed at MHI, Ukrain, has been provided to the team), or indirectly (for instance, issues on image assimilation for meteorology are studied in collaboration with operational centres).

The research activities tackle scientific issues such as:

- Within a family of models (differing by their physical formulations and numerical approximations) which is the optimal model for a given set of observations?

- How to make a forecast (and a better forecast!) by using several models corresponding to different physical formulations? It also raises the question: how should data be assimilated in this context?

- Which observational network should be set up to perform a better forecast, while taking into account additional criteria such as observation cost? What are the optimal location, type and mode of deployment of sensors? How should the trajectories of mobile sensors be operated, while the studied phenomenon is evolving in time? This issue is usually referred to "network design".

- How to assess the quality of the prediction? How do data quality, missing data, data obtained from sub-optimal locations, affect the forecast? How to better include information on uncertainties (of data, of models) within the data assimilation system?

3.2. Environmental image and data processing

Data with image nature, and especially satellite data, represent a huge amount of observations which is up to now largely unexploited by the environmental numerical forecast models. The operational state-of-the-art is mainly the assimilation of satellite data on a pixel basis: each pixel constitutes an independent information, expressed as a more or less complex function of the model's state variables. The challenge is to exploit the structure of the image observation by defining Image Assimilation methods: how to assimilate data with spatial and temporal coherency, such as observations of evolving fronts or eddies? Different issues are considered:

- Rewriting ill-posed image processing problems, usually addressed using numerical regularization techniques, through Image models. The Image Model describes the dynamics of the image sequence and makes it possible to formulate a data assimilation problem, where image observations are assimilated within the Image Model. This approach constitutes a relevant way to solve image processing problems, in which difficult issues such as occlusions or missing data are considered in a natural way. The usual spatial regularization is replaced by the temporal evolution laws for solving the underdetermination issue.

- Definition of Physical Image Models coupling variables image domain and from the forecasting model (in the same spirit than the qualitative Conceptual Models developed by meteorologists to describe specific phenomena and their signature on image data). The assimilation is then performed in two steps: first, in the Physical Image Model to yield "bogus" observations of the forecasting model's state variables, then directly in the forecasting model.

- Learning Image Models from image data. The aim is to define reduced basis, on which projecting the Navier-Stokes equations, to express the dynamics of the image sequence.

- Correcting location of structures from image data. The objective is to define data assimilation methods to modify the position of structures in case of a wrong location in the model representation.

3.3. Software chains for environmental applications

An objective of Clime is to participate in the design and creation of software chains for impact assessment and environmental crisis management. Such software chains bring together static or dynamic databases, data assimilation systems, forecast models, processing methods for environmental data and images, complex visualization tools, scientific workflows, ...

Clime is currently building, in partnership with École des Ponts ParisTech and EDF R&D, such a system for air pollution modeling: Polyphemus (see web site http://cerea.enpc.fr/polyphemus/), whose architecture is specified to satisfy the data requirements (e.g., various raw data natures and sources, data preprocessing) and to support different uses of an air quality model (e.g., forecasting, data assimilation, ensemble runs).

4. Application Domains

4.1. Panorama

The central application of the project-team is atmospheric chemistry, to which a major part of resources are allocated. We develop and maintain the air quality modeling system Polyphemus, which includes several numerical models (Gaussian models, Lagrangian model, two 3D Eulerian models including Polair3D) and their adjoints, and different high level methods: ensemble forecast, sequential and variational data assimilation algorithms. Advanced data assimilation, network design, inverse modeling, ensemble forecast are studied in the context of air chemistry–note that addressing these high level issues requires controlling the full software chain (models and data assimilation algorithms).

The activity on assimilation of satellite data is mainly carried out for meteorology and oceanography. This is addressed in cooperation with external partners who provide the numerical models. Concerning oceanography, the aim is to improve the forecast of ocean circulation, in relation with global warming issues. Concerning meteorology, the focus is on the location of structures related to high-impact weather events (cyclones, convective storms, *etc*.). The underlying researches concern the assimilation of the structured information observed in satellite images.

4.2. Air quality

Air quality modeling implies studying the interactions between meteorology and atmospheric chemistry in the various phases of matter, which leads to the development of highly complex models. The different usages of these models comprise operational forecasting, case studies, impact studies, *etc*, with both societal (e.g., public information on pollution forecast) and economical impacts (e.g., impact studies for dangerous industrial sites). Models lack some appropriate data, for instance better emissions, to perform an accurate forecast and data assimilation techniques are recognized as a key point for the improvement of forecast's quality. These techniques, and notably the variational ones, are progressively surfacing in atmospheric chemistry.

In this context, Clime is interested in various problems, the following being the crucial ones:

- The development of ensemble forecast methods for estimating the quality of the prediction, in relation with the quality of the model and the observations. Sensitivity analysis with respect to the model's parameters so as to identify physical and chemical processes, whose modeling must be improved.

- The development of methodologies for sequential aggregation of ensemble simulations. What ensembles should be generated for that purpose, how spatialized forecasts can be generated with aggregation, how can the different approaches be coupled with data assimilation?

- The definition of second-order data assimilation methods for the design of optimal observation networks. Management of combinations of sensor types and deployment modes. Dynamic management of mobile sensors' trajectories.

- How to estimate the emission rate of an accidental release of a pollutant, using observations and a dispersion model (from the near-field to the continental scale)? How to optimally predict the evolution of a plume? Hence, how to help people in charge of risk evaluation for the population?

- The assimilation of satellite measurements of troposphere chemistry.

The activities of Clime in air quality are supported by the development of the Polyphemus air quality modeling system. This system has a modular design which makes it easier to manage high level applications such as inverse modeling, data assimilation and ensemble forecast.

4.3. Oceanography

The capacity of performing a high quality forecast of the state of the ocean, from the regional to the global scales, is a major requirement of global warming studies. Such a forecast can only be obtained by systematically coupling numerical models and observations (*in situ* and satellite data). In this context, being able to assimilate image structures becomes fundamental. Examples of such structures are:

- apparent motion linked to surface velocity;

- trajectories, obtained either from tracking of features or from integration of the velocity field;

- spatial structures, such as fronts, eddies or filaments.

Image Models for these structures are developed taking into account the underlying physical processes. Image data are assimilated within the Image Models to derive pseudo-observations of the state variables which are further assimilated within the numerical ocean forecast model.

4.4. Meteorology

Meteorological forecasting constitutes a major applicative challenge for Image Assimilation. Although satellite data are operationally assimilated within models, this is mainly done on an independent pixel basis: the observed radiance is linked to the state variable via a radiative transfer model, that plays the role of an observation operator. Indeed, because of their limited spatial and temporal resolutions, numerical weather forecast models fail to exploit image structures, such as precursors of high impact weather:

- cyclogenesis related to the intrusion of dry stratospheric air in the troposphere (a precursor of cyclones);

- convective systems (supercells) leading to heavy winter time storms;

- low-level temperature inversion leading to fog and ice formation, *etc*.

To date, there is no available method for assimilating data which are characterized by a strong coherence in space and time. Meteorologists have developed qualitative Conceptual Models (CMs), for describing the high impact weathers and their signature on images, and tools to detect CMs on image data. The result of this detection is used for correcting the numerical models, for instance by modifying the initialization. The challenge is therefore to develop a methodological framework allowing the assimilation of the detected CMs within numerical forecast models, a very important issue considering the considerable impact of the related meteorological events.

5. Software

5.1. Polyphemus

Participants: Vivien Mallet, Pierre Tran, Irène Korsakissok, Damien Garaud.

Polyphemus (see the web site http://cerea.enpc.fr/polyphemus/) is a modeling system for air quality. As such, it is designed to yield up-to-date simulations in a reliable framework: data assimilation, ensemble forecast and daily forecasts. Its completeness makes it suitable for use in many applications: photochemistry, aerosols, radionuclides, *etc*. It is able to handle simulations from local to continental scales, with several physical models. It is divided into three main parts:

- libraries that gather data processing tools (SeldonData), physical parameterizations (AtmoData) and postprocessing abilities (AtmoPy);

- programs for physical preprocessing and chemistry-transport models (Polair3D, Castor, two Gaussian models, a Lagrangian model);

- drivers on top of the models in order to implement advanced simulation methods such as data assimilation algorithms.

Figure 1 depicts a typical result produced by Polyphemus. Clime is involved in the overall design of the system and in the development of advanced methods in model coupling, data assimilation and ensemble forecast (through drivers and post-processing).

In 2009, two stable versions were released. The main changes are: a Python module for ensemble generation, a complete refactoring of the parallelization (open MP and MPI), the addition of a Lagrangian transport model, the support of WRF, and improvements in the physical models.

5.2. Data assimilation library: Verdandi

Participants: Vivien Mallet, Claire Mouton, Dominique Chapelle [MACS], Philippe Moireau [MACS], Marc Fragu [MACS].
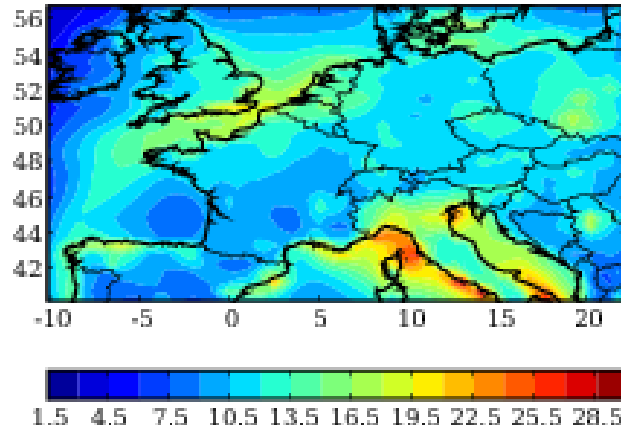
Figure 1. Map of the relative standard deviation (or spread, %) of an ensemble built with Polyphemus (ozone simulations, $\mu g\ m^{-3}$). The standard deviations are averaged over the summer of 2001. They provide an estimation of the simulation uncertainties.

The leading idea is to develop a data assimilation library intended to be generic, at least for high-dimensional systems. Data assimilation methods, developed and used by several teams at INRIA, are generic enough to be written independently of the system to which they are applied. Therefore these methods can be put together in a library aiming at:

- making easier the application of methods to a great number of problems,
- making the developments perennial and sharing them,
- improving the broadcast of data assimilation works.

An object-oriented language (C++) has been chosen for the core of the library. A higher-level interface to Python is automatically built. The design has raised many questions, related to high dimensional scientific computing, the limits of the object contents and their interfaces. The chosen object-oriented design is mainly based on three class hierarchies: the methods, the observation managers and the models. Several base facilities have also been included, for message exchanges between the objects, output saves, logging capabilities, computing with sparse matrices.

The first Verdandi developments offer basic elements allowing to validate the design: a method (optimal interpolation), two linear observation managers and two models (shallow water and clamped bar) have been combined to write the first test programs using Verdandi.

5.3. ENVIAIR

Participants: Isabelle Herlin, Jean-Paul Berroir, Nicolas Mercier.

Enviair is a platform for processing multi-temporal data from the MODIS sensor and extracting land use information. It includes software libraries for:

- Manipulation of satellite images (I/O, management of acquisition dates and geographical coordinates, extraction of temporal data).

- Interpretation of multitemporal information (temporal filtering, computation of temporal features, classification programmes).
- Interpretation of MODIS metadata (pixel quality).

This software has been applied to deforestation monitoring in the high Taquari basin, Brazil (see Figure 2). It is registered commonly by INRIA and Embrapa (Brazilian agricultural research organization).

Figure 2. Left image: classification of low resolution data with the ENVIAIR software (forest in green, pasture in orange, deforestation in pink). Middle and right images: control using two high resolution Landsat images, acquired before and after the low resolution image. The deforestation is clearly visible from the center to the right images.

6. New Results

6.1. Ensemble methods

Due to the great uncertainties that arise in air quality modeling, relying on a single model may not be sufficient. Therefore ensembles of simulations are now considered in a wide range of applications, from uncertainty estimation to operational forecast.

6.1.1. Ensemble forecasting with machine learning algorithms

Participants: Vivien Mallet, Gilles Stoltz [CNRS], Karim Drifi, Édouard Debry [INERIS].

Based on ensemble simulations, improved forecasts can be generated by means of linear combinations of the individual forecasts. A weight is associated to each model, depending on past observations and simulations (Figure 3). New machine learning algorithms (sequential aggregation) were developed and used for this purpose. Most of these provide theoretical bounds on the performance (compared to the optimal constant model combination) and deliver significantly improved forecasts.

The practical performance of the methods which have been developed is very satisfactory. The theoretical bounds are always reached proving that the potential of the ensemble is well exploited. This was checked for large ensembles (dozens of models) as well as for small ensembles (a few models). The methods were successfully applied to forecast ozone, nitrogen dioxide and aerorols in operational mode, on the Prév'air platform (http://www.prevair.org) managed by INERIS. After an operational test from summer 2008 to summer 2009, the methods have been officially introduced in the daily forecasts of Prév'air.

The aggregation methods proved to be efficient on extreme events, but not enough to forecast threshold exceedances: they cannot compensate enough for the poor threshold detection of the individual models. Classification methods, mainly Perceptron, have been studied to address this issue. These methods can slightly improve the forecasts, but further work is needed.
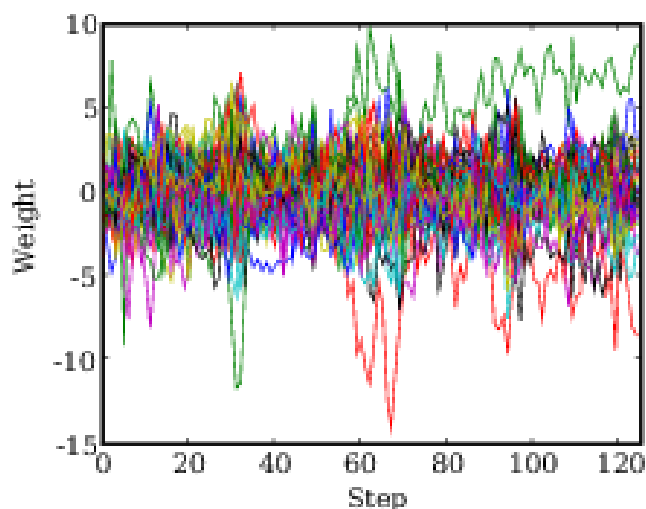
*Figure 3. Weights associated by a machine learning algorithm to the 48 members of an ensemble, against time. The
weights are not constrained (they can be negative).*

6.1.2. *Uncertainty estimation based on multimodel ensembles*
Participants: Damien Garaud, Vivien Mallet.

Air quality forecasts are limited by strong uncertainties especially in the input data and in the physical
formulation of the models. There is a need to estimate these uncertainties for the evaluation of the forecasts, the
production of probabilistic forecasts, and a more accurate estimation of the error covariance matrices required
by data assimilation.

Because a large part of the uncertainty in the forecast originates from uncertainties in the model formula-
tion (primarily the physical parameterizations), a multimodel ensemble seems to be the adequate tool for
uncertainty estimation. A large ensemble with 100 members was generated over year 2001 and analyzed with
criteria like the Brier score. Preliminary work on the calibration of the ensemble was carried out (Figure 4):
the ensemble members were selected so as to optimize the evaluation criteria. This may be formulated as a
combinatorial optimization problem where one searches for an optimal combination of models out of a huge
space of acceptable models.

6.2. Multiscale modeling of air quality
Participants: Vivien Mallet, Irène Korsakissok.

Classical large-scale models in air quality are based on Eulerian approaches. In particular, it is usually assumed
that emissions from the point sources mix immediately within the grid cell, whereas a typical point source
plume (e.g., from a power plant) does not expand to the size of the grid cell for a substantial time period.
Hence, there is a need for a subgrid-scale modeling of the key phenomena (emissions, transport and chemistry).
The plume-in-grid modeling technique, that consists in coupling a local-scale model with an Eulerian model,
has been developed to allow a more accurate representation of sub-grid processes. A sensitivity study was
carried out for passive tracers with ETEX experimental data in order to investigate the influence of the
parameterizations for standard deviations in the puff model, as well as the feedback methods (i.e., the way
the puff is injected in the Eulerian model). Results for chemically reactive plumes have been obtained in the
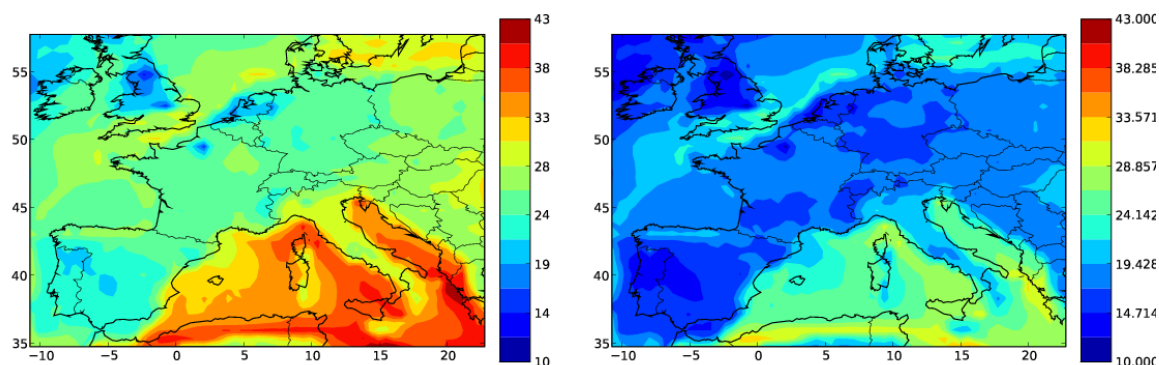Paris area.

Figure 4. Uncertainty estimate (standard deviation) of two ensembles: a raw ensemble on the left, a calibrated ensemble on the right. The uncertainty estimation is significantly reduced after calibration.

Statistical approaches were also studied. Based on a single large-scale model or on an ensemble of large-scale models, statistical downscaling allows to accurately forecast air quality (that is, pollutant concentrations) at observed locations. The methods rely on regressions. The use of an ensemble leads to problems due to the colinearities between the regressors (the models). This issue is addressed with reduction based on principal component analysis or, preferably, with principal fitted components.

6.3. New methods for data assimilation

Since the beginning, the CLIME project has also been focussing on new techniques for data assimilation. Since air quality is prone to non-Gaussian statistics, an expertise has first been on rigorous non-Gaussian approaches, often based on information-theoretical tools (maximum entropy on the mean, relative entropy, second order analysis, etc.). Another expertise is now being developed in multiscale data assimilation, and the mathematical tools required to deal with many space and time scales within data assimilation schemes. It has been made concrete with the launch of the ANR project MSDAG (Multiscale Data Assimilation for Geophysics) in January 2009.

6.3.1. *Towards optimal choices of control space representation for geophysical data assimilation*

Participant: Marc Bocquet.

In geophysical data assimilation, observations shed light on a control parameter space through a model, a statistical prior, and an optimal combination of these sources of information. This control can be a set of discrete parameters, or, more often in geophysics, part of the state vector, which is distributed in space and time. When the control space is continuous, it must be discretised for numerical modeling. This discretisation, called a representation of the distributed parameter space in the framework of this work, is always fixed a priori. The representation of the control space should however be considered a degree of freedom on its own. The goal of this work is to demonstrate that one could optimise it to perform data assimilation in optimal conditions. The optimal representation is then chosen over a large dictionary of adaptive grid representations involving several space and time scales.

First, the importance of the representation choice has been studied through the impact of a change of representation on the posterior analysis of data assimilation and its connection to the reduction of uncertainty. The study stresses that in some circumstances (atmospheric chemistry, in particular) the choice of a proper representation of the control space is essential to set the data assimilation statistical framework properly. A possible mathematical framework has been proposed for multiscale data assimilation. To keep the developments simple, a measure of the reduction of uncertainty is chosen as a very simple optimality criterion. Using this criterion, a cost function is built to select the optimal representation. It is a function of the control space representation itself. A regularisation of this cost function, based on a statistical mechanical analogy, guarantees the existence of a solution. This allows numerical optimisation to be performed on the representation of control space. The formalism has then been successfully applied to the inverse modeling of an accidental release of an atmospheric contaminant at European scale, using real data (see Figure 5).

This is a first contribution from CLIME to the ANR SYSCOMM MSDAG project.
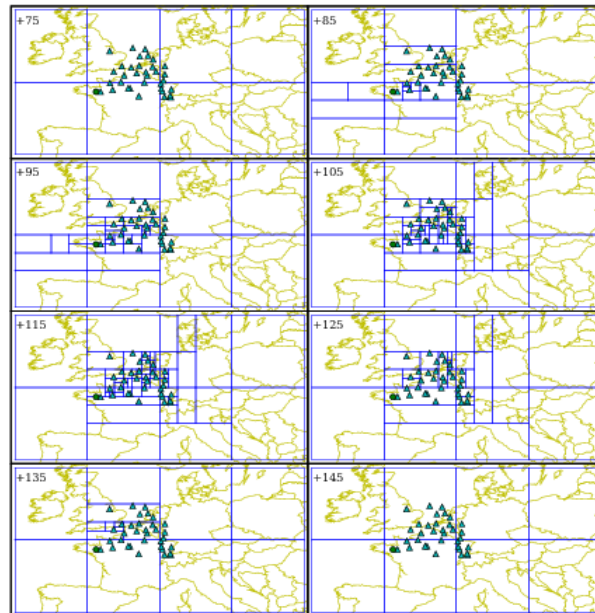

Figure 5. Optimal adaptive grid computed for a subset of observations during the ETEX-I experiment. Since this grid is dynamic - it is a 2D+T object - several time slices of the grid are displayed on the figure.

6.3.2. Modeling non-Gaussianity of background and observational errors by the maximum entropy method

Participants: Carlos A. Pires [Instituto Dom Luis, University of Lisbon, Portugal], Olivier Talagrand [Laboratoire de Météorologie Dynamique], Marc Bocquet.

The Best Linear Unbiased Estimator (BLUE) has widely been used in atmospheric and oceanic data assimilation. However, when the errors from data (observations and background forecasts) have non-Gaussian probability density functions (pdfs), the BLUE differs from the absolute Minimum Variance Unbiased Estimator (MVUE), minimising the mean square a posteriori error. The non-Gaussianity of errors can be due to the inherent statistical skewness and positiveness of some physical observables (e.g., moisture, chemical species) or because of the nonlinearity of the data assimilation models and observation operators acting on Gaussian errors.

Non-Gaussianity of assimilated data errors can be justified from a priori hypotheses or inferred from statistical diagnostics of innovations (observation minus background). Following this rationale, we compute measures of innovation non-Gaussianity, namely its skewness and kurtosis, relating it to: a) the non-Gaussianity of the individual errors themselves, b) the correlation between nonlinear functions of errors, and c) the heteroscedasticity of errors within diagnostic samples. Those relationships impose bounds for skewness and kurtosis of errors which are critically dependent on the error variances, thus leading to a necessary tuning of error variances in order to accomplish consistency with innovations. We evaluate the sub-optimality of the BLUE as compared to the MVUE, in terms of excess of error variance, under the presence of non-Gaussian errors. The error pdfs are obtained by the maximum entropy method constrained by error moments up to fourth order, from which the Bayesian probability density function and the MVUE are computed. The impact is higher for skewed extreme innovations and grows in average with the skewness of data errors, especially if those skewnesses have the same sign. Application has been performed to the quality-accepted ECMWF innovations of brightness temperatures of a set of High Resolution Infrared Sounder (HIRS) channels. In this context, the MVUE has led in some extreme cases to a potential reduction of 20-60% error variance as compared to the BLUE.

6.4. Monitoring network design

6.4.1. *Design of a monitoring network over France in case of a radiological accidental release*
Participants: Marc Bocquet, Olivier Saunier [École des Ponts ParisTech/IRSN].

Launched in March 2006, the network design activity aims at developing new methodologies and applying them to the optimal design of monitoring network for air pollution. Our efforts are dedicated on one hand to the design of atmospheric accidental surveillance networks, and on the other hand to the design of air quality (ozone for instance) monitoring networks. This activity has been supported by the IRSN and Région Île-de-France (R2DS research network). It has been generating discussions with INERIS, ADEME and AIRPARIF.

The Institute of Radiation Protection and Nuclear Safety (France) is planning the setup of an automatic nuclear aerosol monitoring network over the French territory (Descartes network), which complements the Teleray network. Each of the stations will be able to automatically sample the air aerosol content and to provide with activity concentration measurements on several radionuclides. This should help monitor the French and neighbouring countries nuclear power plant park. It would help evaluate the impact of a radiological incident on this park.

After the completion of the first phase (2006 and 2007), the second stage of the study started in March 2008. The resolution has increased from 0.36 °× 0.36 ° to 0.25 °× 0.25°, which doubles the number of potential sites, and hence the complexity of the optimisation problem. Meteorological fields have been generated with MM5 model. New considerations have been taken into account: the inclusion of foreign nuclear power plants, the validation of the optimal network on new cost functions that have not been considered yet, or taking into account the population density as a weighting factor. Moreover, because the Descartes network might be deployed sequentially, we have also considered sub-optimal network design algorithms.

The computational time which was an important issue in the first stage is now a decisive issue because of the resolution increase. In order to accelerate the optimisations, we have developed new reduction techniques for network design optimisation. They are based on the reduction of the database of accidents using ideas derived from principal component analysis. These methods were proven to be very efficient on test cases. They were successfully applied to those new questions that were risen in phase 2 of the Descartes project.

6.4.2. *Reduction of an air quality monitoring network over France*
Participants: Lin Wu, Marc Bocquet.

Ozone is an important air pollutant and observational networks are constructed for its estimation at the ground level. Due to the heterogeneous nature of the ozone field, the way ozone is observed does matter in the estimation of the concentrations. The evaluation of the network is thus of both theoretical and practical interests. In this study, we assess the efficiency of the BDQA (Base de Donnée sur la Qualité de l'Air) network, by investigating a network reduction problem. We examine how well a subset of this network can represent the full network. The performance of a subnetwork is taken to be the root mean square error of the spatial estimations of ozone concentrations over the whole network based on the observations from that subnetwork. Spatial interpolations are conducted for the ozone estimation taking into account the spatial correlations. Several interpolation methods, namely ordinary kriging, simple kriging about means, kriging with means as external drifts, are compared for a reliable estimation. It is found that the statistical information about the means improves significantly the kriging results. We employ a translated exponential model for the spatial correlations. We show that it is necessary to consider the correlation model to be hourly-varying but daily stationary. The network reduction problem is solved using the simulated annealing algorithm. We obtain considerable improvements for the subnetworks with different sizes. In particular, we have shown that keeping only half of the stations allows to reconstruct the hourly values on the missing stations with an average error inferior to the observational error (see Figure 6).
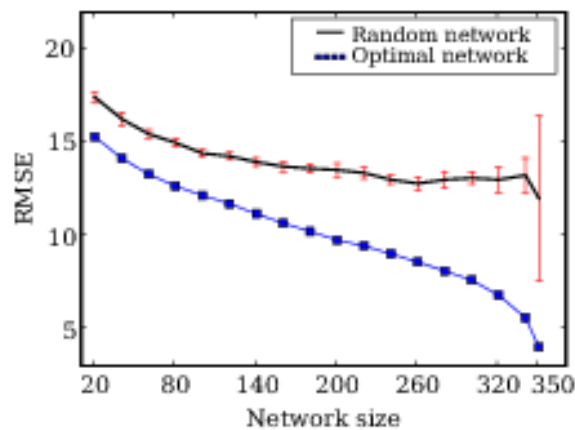

Figure 6. Performances of optimal reduction of the BDAQ network in microgrammes of ozone per cubic meter. These are compared to the performance of random subnetworks of the same size.

6.4.3. Targeting of observations in case of a nuclear accidental release
Participants: Rachid Abida, Marc Bocquet.

In the event of an accidental atmospheric release from a nuclear power plant, high resolution and accurate information on the spread of the radioactive plume around the accident site constitute major key points, acutely required by decision makers in order to evaluate early countermeasure actions and consequences. Therefore, deploying mobile measuring devices constitutes an adequate monitoring strategy that allows to follow the real-time evolution of the radioactive plume. In fact, the collected measurements from the mobile network could be assimilated conjointly with data derived from the fixed monitoring network, so as to enhance knowledge on the state of the radioactive cloud. The targeting design consists in seeking the optimal spatial locations of the mobile stations at a certain time that satisfy some design criterion based on all available previous information. To illustrate how much a targeting strategy could improve the available information on the state of the radioactive plume, we considered an hypothetical accident release occurring at the Bugey power plant

and a sequential data assimilation scheme based on inverse modeling to reconstruct the accident event. This assimilation scheme was coupled with a targeting strategy. The existing surveillance network is used and realistic observational errors are assumed. The targeting scheme leads to a better estimation of the source term as well as the activity concentrations in the domain. The mobile stations tend to be deployed along plume contours, where activity concentration gradients are important. It is shown that the information carried by the targeted observations is very significant, as compared to the information content of fixed observations. A simple test on the impact of model error from meteorology shows that the targeting strategy is still very useful in a more uncertain context.

6.5. Image assimilation

Sequences of images display structures evolving in time. This information is recognized of major interest by, for instance, meteorological forecasters. However, the satellite acquisitions are mostly assimilated in geophysical models on a point-wise basis, discarding the space-time coherence visualized by the evolution of structures. Assimilating images is then becoming of major interest and the problem should be considered in two ways:

- from the model's viewpoint, the problem is to control the location of structures using the observations,
- from the image's viewpoint, a model of the dynamics and structures has to be built from the observations.

In both cases, image information is assimilated within models, raising a number of theoretical and experimental questions.

6.5.1. *Impact of the regularization term*

**Participants:** Etienne Huot, Isabelle Herlin.

The objective is to infer the dynamics from a sequence of satellite images. The application concerns the estimation of surface velocity from Sea Surface Temperature (SST) acquisitions. We define an *Image Model* (*IM*) describing the evolution of the surface temperature and velocity. SST observations are then assimilated in the *IM* by minimizing a cost function, including the measure of the discrepancy between observations and simulations and a regularization term. Two regularization constraints have been compared and tested (see Figure 7): (*i*) the *smoothness* constraint, based on the gradients of the velocity components, and (*ii*) the *second-order* $div - curl$ constraint, based on gradients of irrotational and vorticity components. For quantitative evaluation, synthetic data (computed by an ocean simulation model developed in the MOISE project-team during the ADDISA ANR project) are used. In this context, the *second-order* $div - curl$ regularization is better adapted (4.5% improvement of results compared to the *smoothness* constraint).

6.5.2. *Requirements on the evolution equation*

Participants: Etienne Huot, Isabelle Herlin, Gennady Korotaev [Marine Hydrophysical Institute, Ukraine].

In the context of ocean surface velocity estimation, an *Image Model* (*IM*) is used to express the evolution of the temperature and the dynamics of the velocity. Sea Surface Temperature acquisitions are assimilated in this *IM*, to drive pseudo-observations of velocity, which are further assimilated in the oceanic forecasting model.

Two *Image Models* have been proposed. The *Simple Image Model* (*SIM*) is based on a simplification of the advection-diffusion equation governing the transport of temperature and on the stationarity hypothesis of the velocity field, *i.e.*, it considers that the surface velocity varies much slower than the temperature. Even if this heuristic is often verified, the main drawback is its lack of physical origin to express the dynamics. Hence, an *Extended Image Model* is defined using the same evolution equation for temperature and modeling the velocity through a shallow-water approximation: the evolution of the two components of velocity are linked by the water layer thickness. Results are then compared using first synthetic data, demonstrating the quantitative improvement obtained with the *EIM* (see Figure 8).
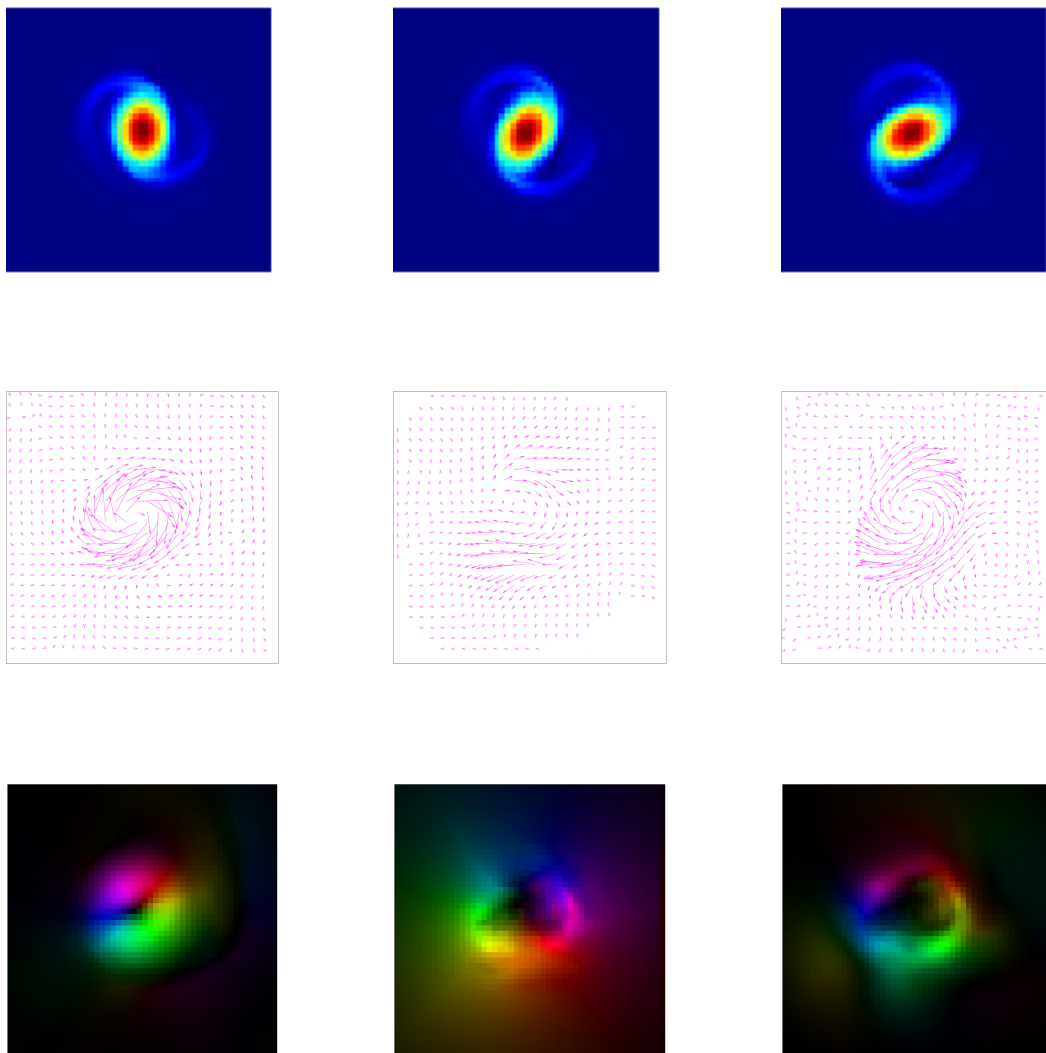
Figure 7. Top. Three simulations of SST data. Middle. Left: velocity computed by the model (ground thruth). Center: with smoothness constraint. Right: second-order $div - curl$ regularization. Bottom. Same fields displayed using an HSV representation.
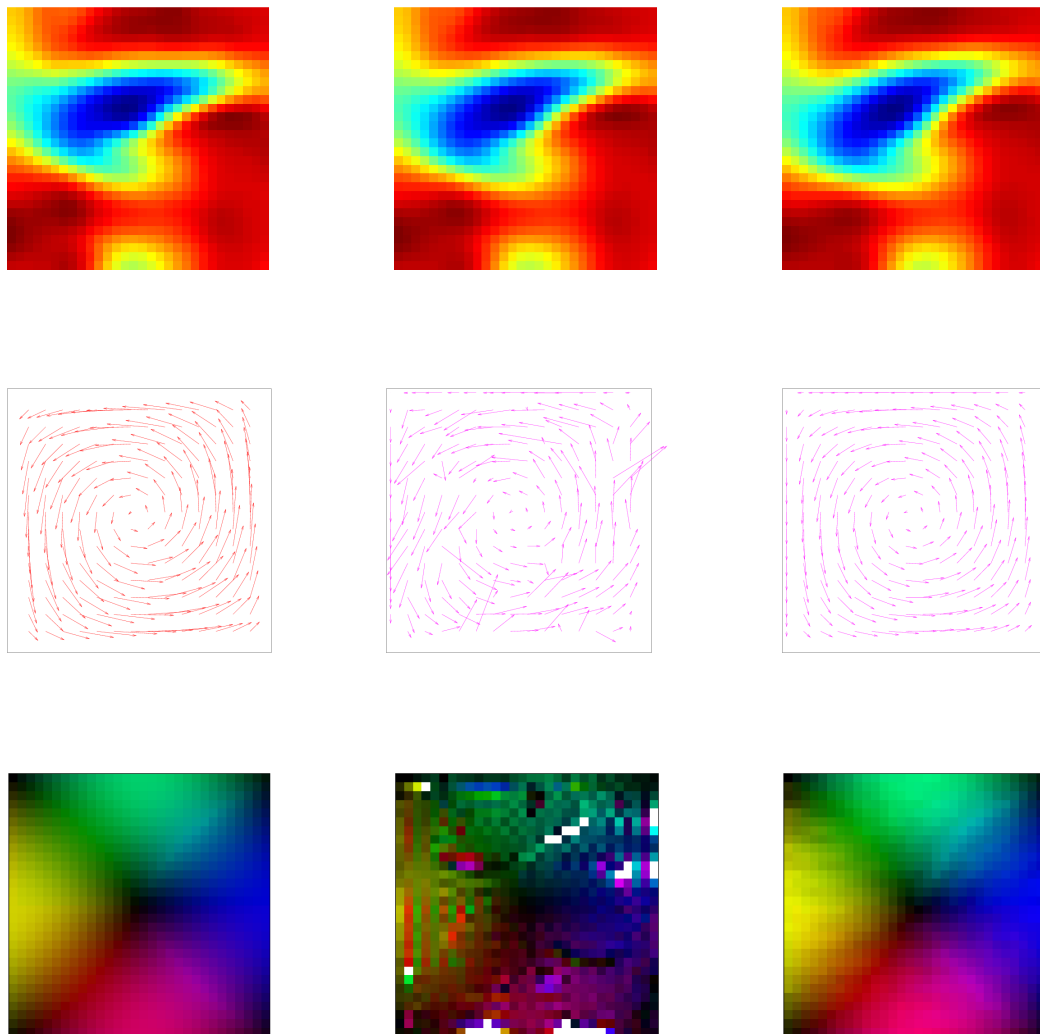
Figure 8. Top: Three simulations of SST data. Middle: ground truth (left), compared to the motion estimated with SIM (center) and with EIM (right). Bottom: same fields displayed using an HSV representation.

6.5.3. Solving ill-posed image processing problem using data assimilation

Participants: Dominique Béréziat [UPMC / LIP6], Isabelle Herlin, Nicolas Mercier, Jean-Paul Berroir.

Most image processing problems are ill-posed in the sense that the *image equation*, modeling the links between the image and the quantity (named the state vector) to be computed is not invertible. A unique solution can however be obtained using a Tikhonov regularization technique. If an *evolution equation*, describing the dynamics of the state vector, is available, it becomes possible to obtain a unique solution, without any regularization, by integrating the evolution equation from the initial condition. Data assimilation offers the mathematical framework to solve simultaneously the image and the evolution equations. We proposed a method to transform, in a generic way, an ill-posed Image Processing problem into a 4D-var formulation. First, state and observation vectors have to be defined. Second, the evolution equation must be exhibited. For some applications, this equation is inferred using physical considerations. However, the dynamics is often unknown and generic models are considered, expressing a temporal regularity of the state vector. Third, model errors associated to the image and evolution equations must be defined. These errors are fully described by their covariance matrices and we studied some generic choices and their impacts on the result. Covariance matrices can also used to process noisy data by discarding the contribution of observations in the computation of the state vector (see Figure 9). Last, an initial condition should be provided. It can be obtained using the traditional approaches: with the image equation and the Tikhonov regularization.

For allowing a generic transformation of an ill-posed image processing problem into a 4D-var formulation, the evolution and observation models are expressed as two operators involved in the evolution and observation equations. These models are discretized and Automatic Differentiation (AD) tools are then used to compute the discretized differentials and adjoints. This enhances the generic aspect of the 4D-var formulation as only observation and evolution models need to be implemented. Moreover, as complex evolution models can lead to unstable numerical schemes, an elegant solution to enhance stability is to split the evolution operator in several simple sub-operators: AD computes differential and adjoint sub-operators.
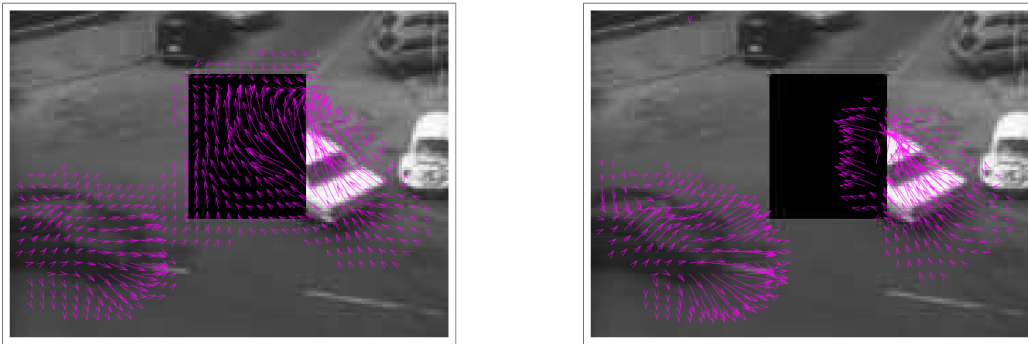

Figure 9. Example of optical flow determination in case of missing data. Left: Data Assimilation. Right: Spatial regularization.

6.5.4. Impact of covariances in 4D-var formulations

Participants: Dominique Béréziat [UPMC / LIP6], Isabelle Herlin, Sahar Syassi.

A general data assimilation algorithm solves, with respect to the state vector, three equations: an evolution equation, an observation equation and an initial condition. Each equation is weighted by a covariance matrix in the functional to be minimized in the variational formulation. The aim of data assimilation is to determine a solution which is a compromise between the observations and the evolution model, given the initial condition. If observations are noisy, they are discarded from the process by imposing high values of the observation error's covariance matrix.

The situation is slightly different in image processing, due to the low confidence in the evolution equation: the image dynamics is usually unknown and only approximated. Consequently, the contribution of that equation in the determination of the state vector has to be lowered. Two problems are then arising.

First, it is no more possible to compute a solution from the observation equation as it is generally ill-posed. The solution is then to add a regularization term, expressed within the observation equation.

Second, the evolution equation errors must be located in time-space. This is achieved by measuring the discrepancy between a solution computed by the evolution equation and a solution computed by the observation equation including the regularization term. This distance is used to specify the covariance associated to the evolution equation error.

6.5.5. *Determining optical flow with large displacement*

Participants: Dominique Béréziat [UPMC / LIP6], Isabelle Herlin.

This addresses video sequences for which the image dynamics is totally unknown. A velocity fields transport of velocity by itself is considered. It is well suited to impose a temporal regularity of the velocity fields. The standard OFCE (Optical Flow Constraint Equation), modeling the image brightness transport by velocity, is applied as observation equation. If large displacements, and therefore high velocities, occur, the OFCE is however no more valid: this PDE is only standing for infinitesimal displacements. The transport of image brightness I by velocity $\mathbf{w}$ between two dates can however be expressed in the following form: $I(\mathbf{x} + \mathbf{w}\delta t, t + \delta t) = I(\mathbf{x})$. This equation is non linear but differentiable. This property is sufficient to apply 4D-var as the algorithm does not need to inverse the observation equation to compute the solution. Successful tests have been performed on synthetic data and video sequences.

6.5.6. *A posteriori guaranteed motion estimation*

Participants: Sergey Zhuk, Isabelle Herlin.

In this study, we focus on the application of the minimax state estimation framework for the motion estimation from an image sequence, using the optical flow equation:

$$\frac{\partial}{\partial t}I + \langle v, \frac{\partial}{\partial x}I\rangle + \mu \triangle I = f(t,x), I(t_0, x) = G(x), x \in P$$

First this equation is rewritten as an observation equation:

$$y(t) = H(t)v(t) + f(t), t_0 \leq t \leq T \tag{1}$$

where $y(t) \mapsto \frac{\partial}{\partial t}I(t,x) + \mu \triangle I(t,x)$, and $H(t)v(t) \mapsto -\langle v(t,x), \frac{\partial}{\partial x}I(t,x)\rangle$ and we consider the evolution equation:

$$\frac{d}{dt}v(t) + L(v) = Bg(t), v(t_0) = v_0, t_0 \leq t \leq T \tag{2}$$

describing the dynamics of the motion field $v(t,x)$. The objective is to estimate v provided that:

$$G(v_0, f, g) := \|Q_0^{\frac{1}{2}}v_0\|^2 + \int_{t_0}^{T} \|R^{\frac{1}{2}}(t)f(t)\|^2 + \left\|Q^{\frac{1}{2}}(t)g(t)\right\|^2 dt \leq 1 \tag{3}$$

where Q_0, Q, R are self-adjoint positive definite bounded linear operators with bounded inverses.

We are looking for the estimation among all functions solving (2) for some v_0, g and verifying:

$$\mathcal{V}(t, v(t)) \leq 1, t_0 \leq t \leq T \tag{4}$$

where $\mathcal{V}$ denotes the value-function:

$$
\begin{aligned}
\mathcal{V}(\tau, v) = \min_{g(\cdot)} \|Q_0^{\frac{1}{2}} v(t_0)\|^2 + \int_{t_0}^{\tau} \left\|R^{\frac{1}{2}}(t)(y(t) - H(t)v(t))\right\|^2 dt \\
+ \int_{t_0}^{\tau} \left\|Q^{\frac{1}{2}}(t)g(t)\right\|^2 dt
\end{aligned} \tag{5}
$$

and $\min$ is taken over all $g(\cdot)$.

We investigate the conditions on the "shape" of the model operator L so that the value function $\mathcal{V}$ verify in a weak sense some Hamilton-Jacobi-Bellman (HJB) equation. We study the problem of solution approximations of the resulting HJB by finite-dimensional HJBs.

6.6. Assimilation of NO_2 columns for air quality forecast

Participants: Xiaoni Wang, Jean-Paul Berroir, Vivien Mallet, Isabelle Herlin, Marc Bocquet.

This study concerns the data assimilation of satellite observations for improving the air quality forecast, performed by the Polyphemus air quality system.

Nitrogen dioxide (NO_2) plays an important role in the tropospheric chemistry and has a direct impact on the public health. A better knowledge and forecast of NO_2 concentration are important to all issues related to air quality. In this research work available satellite data are considered: the Ozone Monitoring Instrument (OMI), aboard NASA Aura satellite, provides NO_2 column data with a good spatial resolution (13 by 24 km^2) and daily global coverage.

First, satellite data have been compared to Polyphemus simulations: the OMI column data and the Polyphemus simulations have both been averaged over November-December 2005 in Europe, demonstrating a good consistency in Spain, Italy and North Europe (see Figure 10), even if model simulations have higher values than satellite observations.

The satellite acquisitions are then assimilated in Polyphemus. The forecast obtained with and without assimilation are compared with ground observations for validation. It is found that assimilation of these satellite data improve the NO_2 forecast, with the RMSE between model results and ground observations reduced after assimilation.

7. Contracts and Grants with Industry

7.1. IRSN

A research contract between CEREA (Clime is simultaneously a team of CEREA and a project-team of INRIA) and IRSN is underway on the topic of network design. Its objective is the optimal construction of a radionuclide monitoring network for detection and diagnosis in case of an accident/event in a French or neighbouring country nuclear power plant.

7.2. INERIS

Clime is partner with INERIS (National Institute for Environmental and Industrial Risks) in a joint cooperation devoted to air quality forecast. This includes research topics in uncertainty estimation, data assimilation and ensemble modeling.
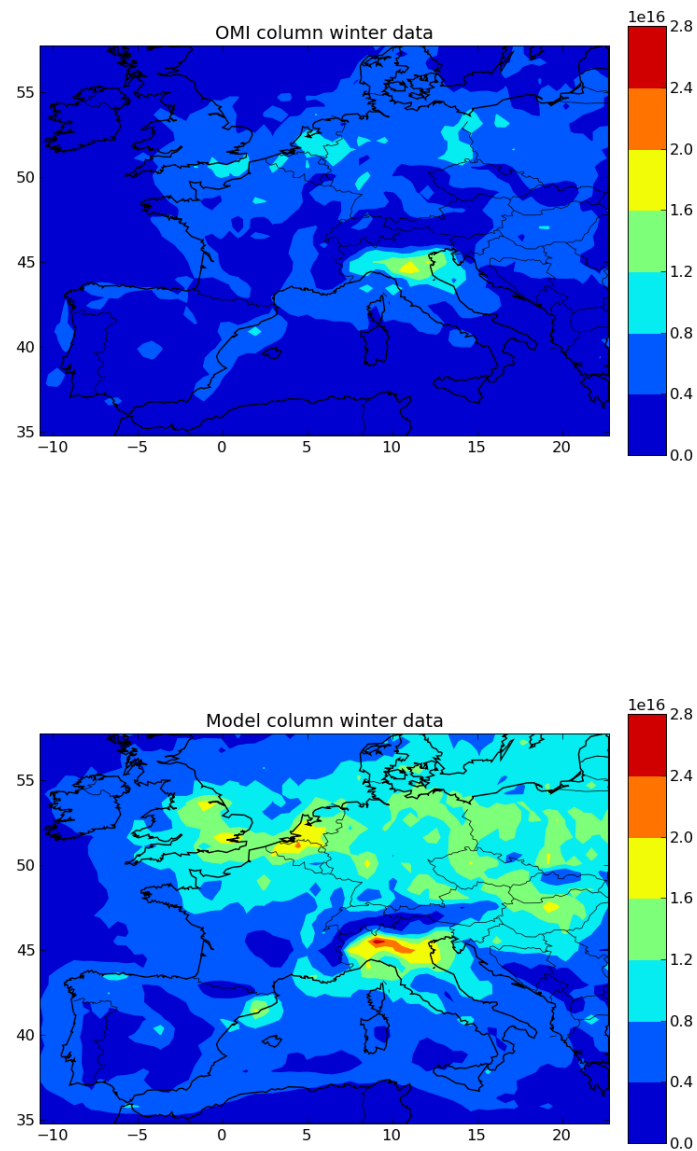
Figure 10. Top: the OMI column data averaged in Europe over November and December of 2005. Down: the Polyphemus simulation of model columns averaged in Europe in the same time period.

Clime also provides support to INERIS in order to operate the Polyphemus system, for ensemble forecasting and uncertainty estimations at local and continental scale.

8. Other Grants and Activities

8.1. National initiatives

- Clime is member of the ADDISA project (ANR, started january 2007) with Moise (Inria Grenoble Rhône-Alpes), LEGI, the CNRM/GAME laboratory of Météo-France and the MIP laboratory of Université Paul Sabatier in Toulouse. This concerns image assimilation applied to meteorology and oceanography.
- Clime is leading ADDISAAF (*Assimilation de Données Distribuées et Images SAtellite pour l'AFrique*) funded by IRD in the Corus program framework, in collaboration with ENIT (Tunisia), the Yaoundé University, and Moise.
- Clime takes part to the ANR project ATLAS ("From Applications to Theory in Learning and Adaptive Statistics"). Clime collaborates with Gilles Stoltz, co-leader of ATLAS, on the application of machine learning to air quality forecasting.
- Clime will take part to the the ANR project IDEA (the project is due to start January 2010) that addresses the propagation of wildland fires. Clime is in charge of the estimation of the uncertainties, based on sensitivity studies and ensemble simulations.
- The three-year project Multiscale Data Assimilation in Geophysics [MSDAG] has been accepted by the ANR SYSCOMM. Fours partners are in the project: CEREA (Clime project-team, Marc Bocquet, PI of the whole project), Fluminance and Moise Project-team, LSCE (Peter Rayner). The preparatory work has led to the definition of a document where an overview of state-of-the-art methodological approaches for multiscale data assimilation is presented. The project has started on January 2009.
- Clime is running the project MIDAR "Inverse modeling of deposition measurements in case of a radiological release", under the framework of the LEFE-ASSIM program of INSU. This includes a cooperation with the Institute for Safety Problems of Nuclear Power Plants (National Academy of Sciences of Ukraine).
- Clime is running an R2DS project "Optimization of Monitoring Networks for Air Quality ", with a grant from Île-de-France region. The aim is to optimally reduce/design a monitoring network for pollutants (ozone in particular).

8.2. European initiatives

- Clime is member of the ERCIM working group "Environmental Modeling". Within this working group, Clime cooperates with FORTH-IACM on remote sensing methods and definition of ontologies for complex applications.

8.3. International initiatives

- Following cooperations with CMM (Chile) on establishing air quality forecast systems and data assimilation capacities in Chile (supported by a research project STIC-AmSud), the Chilean meteorological office (Dirección Meteorológica de Chile) now produces its operational air quality forecasts with Polyphemus. The 3-day forecasts essentially cover Santiago. The forecasts are accessible online in the form of maps, time series and video (http://www.meteochile.cl/modelos.html).
- Clime leads, in cooperation with Moise, the associated team ADAMS (*Advanced Data Assimilation for the Sea*). The partners are the Marine Hydrophysical Institute (Ukrain), the Institute of Numerical Mathematics (Russia) and the Nodia Institute of Geophysics (Georgia).

- An ECO-NET project, ADOMENO, started in 2008, in collaboration with Georgia, Russia and Ukrain. The objectives of ADOMENO are the enhancement of data assimilation techniques by the use of high level data (such as image or Lagrangian data) and advanced assimilation methods. The application domain is the Black Sea circulation.
- A Safeti project (cooperation with South Africa in Information Technologies) is running with F'SATIE, MERAKA Institute (RSA), IRD and ESIEE (France) on the detection, recognition, tracking and characterization of satellite image features for environmental forecast and monitoring, with applications to ocean circulation and fire detection and forecast.

9. Dissemination

9.1. Leadership within scientific community

- Marc Bocquet is co-president of the scientific commitee of the INSU/LEFE action Assimilation.
- Isabelle Herlin is member of the scientific evaluation commitee for the ANR/SYSCOMM program. She is leading the evaluation commitee of international collaborations at INRIA. She is member of the selection commitee of Institut Polytechnique de Grenoble and Université de Créteil.

9.2. Teaching

- Data Assimilation for Geophysics (Master OACOS (ocean, atmosphere, climate and space observation), ENSTA ParisTech/École des Ponts ParisTech): 30 hours (Marc Bocquet, Vivien Mallet).
- Algorithmics: 30 hours, ESIEE Management (Isabelle Herlin).
- Master on nuclear energy: 9 hours (Marc Bocquet, Irène Korsakissok, Vivien Mallet).
- Introduction to chemistry-transport models (Paris VII): 4 hours (Vivien Mallet).
- Air Pollution (École des Ponts ParisTech): 3h00 (Vivien Mallet).

9.3. Conference and workshop committees, invited conferences

- Marc Bocquet:
  - Invited seminar in honor of F.-X. Le Dimet, LJK/MOISE, Grenoble, March 2009.
  - Oral presentation, EGU 2009, Vienna, Austria, April 2009.
  - Invited seminar, CEA DAM, Bruyère-le-Chatel, May 2009.
  - Posters and chairman, 5th WMO data assimilation workshop, Merlbourne, Australia, October 2009.
  - Invited seminar, FMI, Helsinki, Finland, November 2009.
  - Invited seminar, University of Reading, Meteorological department, November 2009.
  - Oral presentation, at a LEFE-ASSIM workshop, Paris, December, 2009.
- Damien Garaud:
  - Oral presentation, ACCENT/GLOREAM Workshop, Brescia, Italy, November, 2009.
- Isabelle Herlin:
  - Oral presentation, Addisa workshop, Toulouse, February 2009.
  - Invited seminar, KNMI, Amsterdam, Netherland, February 2009.
  - Invited seminar, ZAMG, Vienne, Austria, February 2009.
  - Invited seminar, MHI, Sébastopol, Ukrain, August 2009.

- Oral presentation, Adams workshop, Paris, October 2009.
- Vivien Mallet:
  - Invited seminar, University College London, January 2009.
  - Invited seminar, Université d'Orsay, February 2009.
  - Invited seminar, CERMICS, École des ponts ParisTech, March 2009.
  - Oral presentation, "Mathématiques et industries", IHES, April 2009.
  - Oral presentation, EGU 2009, Vienna, Austria, April 2009.
  - Oral presentation, at a LEFE-ASSIM workshop, Paris, December, 2009.

9.4. Visiting scientists

- Pr. Vasudeva Murthy from TIFR-CAM, India: from January 5th to 21th.
- Pablo Saide from Centro Matematico Modeliamento, Santiago of Chile: 2 weeks in March 2009.
- Pr. Valery Agoshkov from Institute of Numerical Mathematics (Russian Academy of Sciences), Russia: from October 25th to November 1st.
- Evgeny Botvinovskiy from Institute of Numerical Mathematics (Russian Academy of Sciences), Russia: from October 25th to November 1st.
- Pr. Demuri Demetrashvili from Mikheil Nodia Institute of Geophysics, Georgia: from October 25th to November 1st.
- Pr. Avtandil Kordzadze from Mikheil Nodia Institute of Geophysics, Georgia: from October 25th to November 1st.
- Pr. Gennady Korotaev from the Marine Hydrophysical Institute of Sebastopol, Ukraine: from October 25th to November 1st.
- Pr. Evgeny Parmuzin from Institute of Numerical Mathematics (Russian Academy of Sciences), Russia: from October 25th to November 1st.
- Ievgen Plotnikov from the Marine Hydrophysical Institute of Sebastopol, Ukraine: from October 25th to November 1st.
- Pr. Victor Shutyaev from Institute of Numerical Mathematics (Russian Academy of Sciences), Russia: from October 25th to November 1st.

10. Bibliography

Major publications by the team in recent years

[1] M. BOCQUET. *Reconstruction of an atmospheric tracer source using the principle of maximum entropy*, in "Journal of the Royal Meteorological Society", vol. 131, part B, 2005, 2191.

[2] M. BOCQUET. *High resolution reconstruction of a tracer dispersion event: application to ETEX*, in "Q. J. R. Meteorol. Soc.", vol. 133, 2007, p. 1013–1026.

[3] D. BÉRÉZIAT, I. HERLIN. *Solving Ill-posed Problems Using Data Assimilation. Application to optical flow estimation*, in "VISAPP'09, Lisboa (Portugal)", Feburary 2009.

[4] I. HERLIN, E. HUOT, J.-P. BERROIR, F. LE DIMET, G. KOROTAEV. *Estimation of a motion field on satellite images from a simplified ocean circulation model*, in "ICIP International Conference on Image Processing, Atlanta, USA", October, 8-11 2006.

[5] I. HERLIN, F. LE DIMET, E. HUOT, J.-P. BERROIR. *e-Environement: progress and challenges*, in "Coupling models and data: which possibilities for remotely-sensed images?", P. PRASTACOS, J. MURILLO, J. L. DÍAZ DE LEÓN, U. CORTÉS (editors), Instituto Politécnico Nacional, Mexico, 2004, p. 365–383.

[6] M. KRYSTA, M. BOCQUET. *Source reconstruction of an accidental radionuclide release at European scale*, in "Quarterly Journal of the Royal Meteorological Society", vol. 133, 2007, p. 529-544.

[7] V. MALLET, B. SPORTISSE. *Uncertainty in chemistry-transport model due to physical parameterizations and numerical approximations: An ensemble approach applied to ozone modeling*, in "Journal of Geophysical Research", vol. 111, n^o D1, 2006.

[8] V. MALLET, G. STOLTZ, B. MAURICETTE. *Ozone ensemble forecast with machine learning algorithms*, in "Journal of Geophysical Research", vol. 114, n^o D05307, March 2009, http://dx.doi.org/doi:10.1029/2008JD009978.

[9] B. SPORTISSE. *Pollution atmosphérique: des processus à la modélisation*, Springer France, 2007, 350 pages.

[10] L. WU, V. MALLET, M. BOCQUET, B. SPORTISSE. *A Comparison Study of Data Assimilation Algorithms for Ozone Forecasts*, in "J. Geophys. Res.", vol. 113, n^o D20310, 2008, http://dx.doi.org/10.1029/2008JD009991.

Year Publications

Articles in International Peer-Reviewed Journal

[11] R. ABIDA, M. BOCQUET. *Targeting of observations for accidental atmospheric release monitoring*, in "Atmos. Env.", vol. 43, 2009, p. 6312–6327.

[12] M. BOCQUET. *Towards optimal choices of control space representation for geophysical data assimilation*, in "Month. Wea. Rev.", vol. 137, 2009, p. 2331–2348.

[13] D. BÉRÉZIAT, I. HERLIN. *Solving ill-posed Image Processing problems using Data Assimilation*, in "Numerical Algorithms", 2009, accepted.

[14] D. GARAUD, V. MALLET. *Automatic generation of large ensembles for air quality forecasting using the Polyphemus system*, in "Geoscientific Model Development Discussion", 2009, p. 889–933, http://www.geosci-model-dev-discuss.net/2:889/2009/gmdd-2-889-2009.html.

[15] I. KORSAKISSOK, V. MALLET. *Comparative study of Gaussian dispersion formulae within the Polyphemus platform: evaluation with Prairie Grass and Kincaid experiments*, in "Journal of Applied Meteorology and Climatology", 2009, in press.

[16] I. KORSAKISSOK, V. MALLET. *Subgrid-scale treatment for major point sources in an Eulerian model: a sensitivity study on the ETEX and Chernobyl cases*, in "Journal of Geophysical Research", 2009, accepted.

[17] I. KORSAKISSOK, V. MALLET, B. SPORTISSE. *Assessment of the detection abilities of monitoring networks for passive tracers at local and regional scales*, in "Atmospheric Environment", 2009, accepted.

[18] V. MALLET, D. KEYES, F. FENDELL. *Modeling Wildland Fire Propagation with Level Set Methods*, in "Computers & Mathematics with Applications", vol. 57, n° 7, 2009, p. 1,089–1,101, http://dx.doi.org/doi:10.1016/j.camwa.2008.10.089US.

[19] V. MALLET, G. STOLTZ, B. MAURICETTE. *Ozone ensemble forecast with machine learning algorithms*, in "Journal of Geophysical Research", vol. 114, n^o D05307, March 2009, http://dx.doi.org/doi:10.1029/2008JD009978.

[20] O. SAUNIER, M. BOCQUET, A. MATHIEU, O. ISNAR. *Model reduction via principal component truncation for the optimal design of atmospheric monitoring networks*, in "Atmos. Env.", vol. 43, 2009, p. 4940–4950.

[21] E. TOMAI, I. HERLIN, J.-P. BERROIR, P. PRASTOCOS. *Ontology-based documentation of land degradation assessment from satellite images*, in "International Journal of Remote Sensing", vol. 30, n° 13-14, July 28 2009, p. 3315–3330, Taylor & Francis (Publisher), doi=10.1080/01431160802558709 GR .

[22] M. TOMBETTE, V. MALLET, B. SPORTISSE. PM_{10} *data assimilation over Europe with the optimal interpolation method*, in "Atmospheric Chemistry and Physics", vol. 9, n° 1, 2009, p. 57–70.

International Peer-Reviewed Conference/Proceedings

[23] D. BÉRÉZIAT, I. HERLIN. *Solving Ill-posed Problems Using Data Assimilation. Application to optical flow estimation*, in "VISAPP'09, Lisboa (Portugal)", Febuary 2009.

[24] É. DEBRY, V. MALLET, F. MELEUX, B. BESSAGNET, L. ROUÏL. *Improving predictions and threshold detection with ensemble modelling in France*, in "7th International Conference on Air Quality – Science and Application", 2009.